On the Convergence of the Modified Overrelaxation Method

M. Madalena Martins*
Department of Mathematics
University of Coimbra
Portugal

Submitted by Hans Schneider

ABSTRACT

The modified overrelaxation (MSOR) method is applied to a linear system $Ax = b$, where A has property A. We get bounds for the spectral radius of the iteration matrix of this method, and we achieve convergence conditions for the MSOR method when A is strictly diagonally dominant. We extend our conclusions to another kind of matrices— H, L, M or Stieltjes. In the last section we use the vectorial norms, getting convergence conditions for the MSOR method, when A is a block-H matrix. We also generalize a theorem of Robert's for this kind of matrices.

1. INTRODUCTION

Let us denote by $\mathbb{C}^{n,n}$ the class of all complex matrices $A=[a_{i,j}]$ of order $n(n \geqslant 2)$ with $a_{ij} \in \mathbb{C}$, $i, j \in \langle n \rangle$; by A^T the transpose of the matrix A; and by $\mathbb{C}^n$ the complex n-dimensional vector space of column vectors $b=[b_1, b_2, \dots, b_n]^T$, with $b_i \in \mathbb{C}$, $i \in \langle n \rangle$ ($\langle n \rangle$ is the set $\{1,2,\dots,n\}$). The corresponding symbols when all the elements involved are real are respectively $\mathbb{R}^{n,n}$ and $\mathbb{R}^n$.

Definition 1. A matrix $A \in \mathbb{C}^{n,n}$ has property A if and only if A is diagonal or there exists a permutation matrix P such that $P^{-1}AP$ has the form

$$A' = P^{-1}AP = \begin{bmatrix} D_1 & T \\ S & D_2 \end{bmatrix} \tag{1.1}$$

where D_1 and D_2 are square diagonal matrices.

---

*Centro de Matemática da Universidade de Coimbra (CMUC). Supported by Instituto Nacional de Investigação Científica.

52 Vanderbilt Ave., New York, NY 10017 0024-3795/86/$3.50

Let us consider a system of n linear equations:

$$Ax = b, \tag{1.2}$$

where $A \in \mathbf{C}^{n,n}$, $x, b \in \mathbf{C}^n$ with b known and x unknown. We assume that A of (1.2) satisfies property A with D_1 and D_2 of (1.1) nonsingular diagonal matrices of order $k, n-k$, respectively . Thus, to get the solution of (1.2) is equivalent to solving

$$\begin{aligned} x_1 &= Mx_2 + q_1, \\ x_2 &= Nx_1 + q_2 \end{aligned} \tag{1.3}$$

with $M = -D_1^{-1}TN = -D_2^{-1}S$, $q_1 = -D_1^{-1}b_1$, $q_2 = -D_2^{-1}b_2$.

Let us remark that the partition used for x and b is in accord with the splitting used for A.

If we apply to (1.3) the SOR method with the parameters w, w', we obtain the modified SOR method (MSOR)

$$x^{(i+1)} = L_{w,w'}x^{(i)} + K_{w,w'}, \qquad i = 0,1,2, \tag{1.4}$$

where

$$L_{w,w'} = \begin{bmatrix} I_1 & 0 \\ -w'N & I_2 \end{bmatrix}^{-1} \begin{bmatrix} (1-w)I_1 & wM \\ 0 & (1-w')I_2 \end{bmatrix}$$

and

$$K_{w,w'} = \begin{bmatrix} wq_1 \\ ww'Nq_1 + w'q_2 \end{bmatrix}$$

(see Young [11]). We can write $L_{w,w'}$ and $K_{w,w'}$ slightly modified:

$$L_{w,w'} = \left[\begin{pmatrix} I_1 & 0 \\ 0 & I_2 \end{pmatrix} - w' \begin{pmatrix} 0 & 0 \\ N & 0 \end{pmatrix} \right]^{-1} \left[(1-w)I + \begin{pmatrix} 0 & wM \\ 0 & (w-w')I_2 \end{pmatrix} \right]$$

or

$$L_{w,w'} = (I - w'Q)^{-1}[(1-w)I + wB + (w-w')R] \tag{1.5}$$

or

$$L_{w,w'} = (I + w'Q)[(1-w)I + wB + (w-w')R], \tag{1.6}$$

where

$$Q = \begin{bmatrix} 0 & 0 \\ N & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 & M \\ 0 & 0 \end{bmatrix}, \qquad R = \begin{bmatrix} 0 & 0 \\ 0 & I_2 \end{bmatrix}; \tag{1.7}$$

and

$$K_{w,w'} = w \begin{bmatrix} I_1 & 0 \\ -w'N & I_2 \end{bmatrix}^{-1} \begin{bmatrix} q_1 \\ \dfrac{w'}{w} q_2 \end{bmatrix}$$

or

$$K_{w,w'} = w(I - w'Q)^{-1} d \tag{1.8a}$$

or

$$K_{ww'} = w(I + w'Q)d, \tag{1.8b}$$

with

$$d = \begin{bmatrix} q_1 \\ \dfrac{w'}{w} q_2 \end{bmatrix}.$$

The relaxation factor can change from iteration to iteration and differ for the "red" points (corresponding to D_1) and the "black" points (corresponding to D_2).

If all the w's are equal we have the ordinary SOR method, whose iteration matrix we denote by $\mathscr{L}_w^m$.

In Chapter 10 of [11], Young shows that as far as the spectral radius is concerned the best choice is to have all w's equal. However, as far as a certain norm is concerned, the cyclic Chebyshev semi-iterative method (CCSI method), where the w's do vary, is better.

Clearly the MSOR is better than the Jacobi method and Gauss-Seidel method, since the SOR method is better than these methods.

Next we will show the interest of using this method.

DEFINITION 2. Given an $n \times n$ matrix A, we define its *L-norm* by

$$\|A\|_L = \|LAL^{-1}\|,$$

where $\|\cdot\|$, denotes the spectral norm.

As we know (see [11, Chapter 10]), several other iterative methods—Sheldon's method, the modified Sheldon method, and the cyclic Chebyshev semi-iterative method—are only variants of the MSOR method (with variable parameters). Thus, with

$$w_b = \frac{2}{1+\left(1+\bar{\mu}^2\right)^{1/2}} \quad \text{and} \quad \bar{\mu} = \rho(B^*),$$

where B^* is the iteration matrix of the Jacobi method, we have:

(1) *The Sheldon method* [7], whose iteration matrix $\mathscr{H}_m$ has the relaxation parameters

$$w_1 = w_1' = 1, \qquad w_2 = w_2' = w_3 = w_3' = \cdots = w_b.$$

(2) *The modified Sheldon method* [1], whose iteration matrix K_m has the relaxation parameters

$$w_1 = 1, \qquad w_1' = w_2 = w_2' = \cdots = w_b.$$

(3) *The cyclic Chebyshev semi-iterative method* [1], whose iteration matrix C_m has the relaxation parameters

$$w_1 = 1, \qquad\qquad w_1' = \frac{2}{2-\bar{\mu}^2},$$

$$w_k = \left(1 - \tfrac{1}{4}w_{k-1}'\bar{\mu}^2\right)^{-1}, \qquad w_k' = \left(1 - \tfrac{1}{4}w_k\bar{\mu}^2\right)^{-1}, \qquad k = 2,3,\ldots.$$

Referring to Theorems 7.4.1, 7.6.1, 10.4.1, 10.5.1, and 10.6.1 of [11], we can see that for the norm $D^{1/2}$ the cyclic Chebyshev semi-iterative method and Sheldon's method are better than the well-known SOR method.

Looking at the formulas for the $F_i(m)$, $i=1,\dots,4$, given by the above-mentioned theorems, we can also prove (see [1]) that

$$F_1(m) \geqslant F_3(m) \qquad \text{for} \quad m \geqslant 2,$$

$$F_3(m) \leqslant F_4(m) \leqslant F_2(m) \qquad \text{for} \quad m \geqslant 1.$$

Thus the CCSI method is the best method with respect to the norm $D^{1/2}$, followed by the modified Sheldon method, Sheldon's method, and finally the SOR method.

For the norms $A^{1/2}$, as is shown by [12], we have

$$H_2(m) \geqslant H_4(m)$$

for $m \geqslant \frac{3}{2}$ and $r \leqslant 0.840$. Thus the CCSI method is the best method with respect to the norms $D^{1/2}$ and $A^{1/2}$. For the norm $A^{1/2}$ the SOR method is not quite as good as the two Sheldon methods; however, the difference is much less for r near unity than was the case with $D^{1/2}$.

These conclusions can be confirmed with a study based on the theoretical number of iterations (see [11, pp. 331–340]).

2. CONVERGENCE CONDITIONS

THEOREM 1. *The spectral radius $\rho(L_{w,w'})$ of $L_{w,w'}$ satisfies*

$$\min_i \frac{|1-w|-|w|b_i-|w-w'|r_i}{1+|w'|q_i}$$
$$\leqslant \rho(L_{w,w'}) \leqslant \max_i \frac{|w-w'|r_i+|w|b_i+|1-w|}{1-|w'|q_i}, \qquad i\in\langle n\rangle, \quad (2.1)$$

if $|w'|<1/q_i$, where q_i, r_i, b_i are respectively the sums of the absolute values of the elements of the ith row of the matrices Q, R, and B.

Proof. The eigenvalues of $L_{w,w'}$ are given by the roots of

$$\det(\lambda I - L_{w,w'}) = 0.$$

The roots of these equations are the roots of

$$\det P = 0$$

with

$$P = I - \frac{\lambda w'}{\lambda - 1 + w} Q - \frac{w}{\lambda - 1 + w} B - \frac{w - w'}{\lambda - 1 + w} R \tag{2.2}$$

if $\lambda - 1 + w \neq 0$.

Let us suppose that some eigenvalue λ of $L_{w,w'}$ satisfies

$$|\lambda| > \frac{|w - w'|r_i + |w|b_i + |1 - w|}{1 - |w'|q_i}$$

with $|w'| < 1/q_i$, $i \in \langle n \rangle$, or—which is equivalent—

$$|\lambda|(1 - |w'|q_i) > |w - w'|r_i + |w|b_i + |1 - w|, \qquad i \in \langle n \rangle,$$

or

$$|\lambda| - |1 - w| > |w - w'|r_i + |\lambda|\,|w'|q_i + wb_i, \qquad i \in \langle n \rangle.$$

This last relation implies

$$|\lambda - 1 + w| > |w - w'|r_i + |\lambda|\,|w'|q_i + |w|b_i, \qquad i \in \langle n \rangle, \tag{2.3}$$

which means that P is a strictly diagonally dominant matrix, and therefore nonsingular, which is a contradiction.

Thus the values of λ verifying (2.3) cannot be eigenvalues of $L_{w,w'}$, then we must have

$$\rho(L_{w,w'}) \leqslant \max_i \frac{|w - w'|r_i + |w|b_i + |1 - w|}{1 - |w'|q_i}, \qquad i \in \langle n \rangle,$$

with $|w'| < 1/q_i$. In the same way, if an eigenvalue λ of $L_{w,w'}$ satisfies

$$|\lambda| < \frac{|1 - w| - |w|b_i - |w - w'|r_i}{1 + |w'|q_i}, \qquad i \in \langle n \rangle,$$

we have

$$|1 - w| - |\lambda| > |w'|q_i|\lambda| + |w|b_i + |w - w'|r_i, \qquad i \in \langle n \rangle,$$

which implies the relation (2.3). Thus we can conclude that we must have

$$\rho(L_{w,w'}) \geqslant \min_i \frac{|1-w| - |w|b_i - |w-w'|r_i}{1+|w'|q_i}, \qquad i \in \langle n \rangle. \qquad \blacksquare$$

THEOREM 2. *If A of (1.2) is strictly diagonally dominant, then $\rho(L_{w,w'}) < 1$ if*

(i) $$0 < w' < \min_i \frac{b_i}{1-q_i} \text{ and } w' < w < \min_i \frac{w'(1-q_i)}{b_i}, \qquad i \in \langle n \rangle,$$

or

(ii) $$\max \frac{b_i}{1-q_i} < w' < \frac{2}{1+\max_i (b_i + q_i)}$$

and

$$1 \leqslant w < \min_i \frac{2+w'(1-q_i)}{2+b_i}, \qquad i \in \langle n \rangle,$$

with $w > w'$, or if

(iii) $$0 < w' < \min_i \frac{2-b_i}{1+q_i} \text{ and } \max_i \frac{w'(1+q_i)}{2-b_i} < w \leqslant w', \qquad i \in \langle n \rangle,$$

or

(iv) $$\frac{2}{1+\min_i (q_i + b_i)} < w' < \min_i \frac{2-b_i}{1+q_i}$$

and

$$1 \leqslant w < \min_i \frac{2-w'(1+q_i)}{b_i}, \qquad i \in \langle n \rangle,$$

with $w \leqslant w'$.

Proof. By the last theorem the MSOR method is convergent if

$$|w - w'|r_i + |w|b_i + |1 - w| + |w'|q_i < 1$$

or

$$|w - w'| + |w|b_i + |1 - w| + |w'|q_i < 1.$$

(i): We take $w > w'$, and we define the function

$$f(\delta) = \delta - w' + \delta b_i + |1 - \delta| + w'q_i.$$

We can see that $f(\delta)$ is an increasing function for $0 < \delta < 1$ and $f(0) = 1 - w'(1 - q_i) < 1$ with $w' > 0$. Thus, $f(\delta)$ will be less than 1 for $\delta < w'(1 - q_i)/b_i$. As $\delta < 1$, we must have $0 < w' < b_i/(1 - q_i)$.

(ii): If $\delta \geqslant 1$, then $f(\delta)$ is an increasing function and will be less than 1 when

$$\delta < \frac{2 + w'(1 - q_i)}{2 + b_i}, \qquad i \in \langle n \rangle.$$

As $\delta \geqslant 1$ and $w > w'$, we have

$$w' > \frac{b_i}{1 - q_i} \quad \text{and} \quad \frac{2 + w'(1 - q_i)}{2 + b_i} > w',$$

or equivalently $w' < 2/(1 + b_i + q_i)$.

Thus we have obtained the conditions (i) and (ii) for the convergence of the MSOR method.

(iii): If we take $w < w'$, we have

$$f(\delta) = w' - \delta + \delta b_i + |1 - \delta| + w'q_i,$$

and for $0 < \delta < 1$, $f(\delta)$ is a decreasing function for $b_i < 2$, and $f(\delta) < 1$ if $\delta > w'(1 + q_i)/(2 - b_i)$. As $\delta < 1$, we must have $w' < (2 - b_i)/(1 + q_i)$.

(iv): If $\delta \geqslant 1$, then $f(\delta)$ is an increasing function and $f(\delta) < 1$ if $\delta < [2 - w'(1 + q_i)]/b_i$. With $\delta \geqslant 1$, we have $w' < (2 - b_i)/(1 + q_i)$, $i = 1, \ldots, n$. As $w \leqslant w'$, we have $[2 - w'(1 + q_i)]/b_i < w'$, or equivalently $w' > 2/(1 + b_i + q_i)$.

Let us remark that for proving (iii) and (iv) we must have $b_i < 2$.

In this theorem the matrix $A = I - Q - B$ is strictly diagonally dominant, so the condition $b_i < 2$ is verified. ■

3. GENERALIZED DIAGONAL DOMINANCE AND ITS CONNECTIONS WITH THE MSOR METHOD

DEFINITION 3. A scaling by rows of a matrix A is a matrix DA where D is a diagonal nonsingular matrix. A scaling by columns of a matrix A is a matrix AD' where D' is a diagonal, nonsingular matrix.

DEFINITION 4. A matrix A is generalized diagonally dominant by rows (columns) if there is a scaling on columns (rows) of A by nonzero multipliers such that the obtained matrix A is strictly diagonally dominant by rows (columns).

LEMMA 1. *If $A = I + L + U$ is an (n, n) matrix, then A is generalized diagonally dominant by rows (columns) if and only if there are positive vectors v, v' such that*

$$(I - |L| - |U|)v = v',$$

$$(I - |L| - |U|)^T v = v'.$$

Proof. See [2]. ■

In the following theorems we are going to achieve conditions for convergence of the MSOR method when A of (1.2) is an L, H, M, or Stieltjes matrix. Thus we give the following definitions:

DEFINITION 5. A matrix $A \in \mathbb{R}^{n,n}$ is an L-matrix if

$$a_{ii} > 0, \qquad i = 1,2,\dots,n, \tag{3.1}$$

$$a_{ij} \leqslant 0, \qquad i \neq j, \quad i,j = 1,2,\dots,n \tag{3.2}$$

DEFINITION 6. A matrix $A \in \mathbb{R}^{n,n}$ is a Stieltjes matrix if A is positive definite and if (3.2) holds.

DEFINITION 7. A matrix $A \in \mathbb{R}^{n,n}$ is an M-matrix if (3.2) holds and $A^{-1} \geqslant 0$.

DEFINITION 8. An $n \times n$ matrix A is an H-matrix if the comparison matrix $M(A)$ defined by $m_{ii} = |a_{ii}|$, $1 \leqslant i \leqslant n$, $m_{ij} = -|a_{ij}|$, $1 \leqslant i, j \leqslant n$, $i \neq j$, is an M-matrix.

THEOREM 3. *If $A = I - Q - B$, with Q and B defined by* (1.7), *is an irreducible L-matrix, then the MSOR method is convergent for $0 < w \leqslant w' < 1$ if and only if A satisfies generalized diagonal dominance by rows.*

Proof. The iteration of the MSOR method is

$$L_{w,w'} = (I + w'Q)[(1-w)I + wB + (w - w')R],$$

or

$$L_{w,w'} = (1-w)I + wB + w'(1-w)Q + \cdots .$$

As A is irreducible, the matrix $L_{w,w'}$ is also irreducible, and by the Perron-Frobenius theorem we can say that $L_{w,w'}$ has an eigenvalue $\lambda > 0$ equal to the spectral radius, and to this eigenvalue there corresponds a positive eigenvector x:

$$L_{w,w'}x = \lambda x,$$

or

$$[(1-w)I + wB + (w-w')R]x = \lambda(I - w'Q)x.$$

Then

$$w\left(I - B - \lambda\frac{w'}{w}Q\right)x = (1-\lambda)x + (w - w')Rx.$$

If the MSOR method is convergent, then taking $\lambda = w/w' < 1$, we have

$$w(I - B - Q)x = \left(1 - \frac{w}{w'}\right)x + (w - w')Rx \tag{3.3}$$

It is evident that the first k components of the vector $(1 - w/w')x + (w - w')Rx$ are positive. The last $n - k$ components will be positive, because

$$1 - \frac{w}{w'} + w - w' > 0 \quad \text{and} \quad x > 0$$

As A is an L-matrix we can write (3.3) as follows:

$$w(I - |B| - |Q|)x = Z$$

with

$$Z = \left(1 - \frac{w}{w'}\right)x + (w - w')Rx > 0.$$

Then from Lemma 1, we see that A satisfies generalized diagonal dominance by rows.

The sufficient condition is evident. ∎

THEOREM 4. *If $A = I - Q - B$ is an (n, n) matrix with Q and B given by* (1.6), *and $\hat{A}$ is a matrix obtained from A by a scaling by rows and columns, then the iteration matrices for the MSOR method for A and $\hat{A}$ have the same eigenvalues.*

Proof. We have seen that the iteration matrix of the MSOR method is

$$L_{w,w'} = (I - w'Q)^{-1}[(1 - w)I + wB + (w - w')R].$$

Let us denote by $\hat{A} = \hat{D}_1 A \hat{D}_2$, where $\hat{D}_1$ and $\hat{D}_2$ are nonsingular diagonal (n, n) matrices.

If $A = I - Q - B$, we have

$$\hat{I} = \hat{D}_1 I \hat{D}_2, \qquad \hat{Q} = \hat{D}_1 Q \hat{D}_2, \qquad \hat{B} = \hat{D}_1 B \hat{D}_2,$$

which are respectively the diagonal, strictly lower triangular, and strictly upper triangular part of $\hat{A}$. Then the iteration matrix of the MSOR method for $\hat{A}$ is

$$\begin{aligned}\hat{L}_{w,w'} &= (\hat{I} - w'\hat{Q})^{-1}[(1 - w)\hat{I} + w\hat{B} + (w - w')\hat{R}]\\ &= [\hat{D}_1(I - w'Q)\hat{D}_2]^{-1}[(1 - w)\hat{D}_1 I \hat{D}_2 + w\hat{D}_1 B \hat{D}_2 + (w - w')\hat{D}_1 R \hat{D}_2]\\ &= \hat{D}_2^{-1} L_{w,w'} \hat{D}_2,\end{aligned}$$

where $\hat{R} = \hat{D}_1 R \hat{D}_2$. Thus, $L_{w,w'}$ and $\hat{L}_{w,w'}$ have the same eigenvalues. ∎

COROLLARY 1. *The MSOR method is convergent for A if and only if it is convergent for* $\hat{A}$, *and the rate of convergence is the same.*

COROLLARY 2. *If* $\hat{A}$ *obtained from A is strictly diagonally dominant, the MSOR method is convergent for*

$$0 < w' < \min_i \frac{\hat{b}_i}{1-\hat{q}_i} \text{ and } w' < w < \min_i \frac{w'(1-\hat{q}_i)}{\hat{b}_i}, \qquad i \in \langle n \rangle, \quad (3.4)$$

or

$$\max_i \frac{\hat{b}_i}{1-\hat{q}_i} < w' < \frac{2}{1+\max_i(\hat{b}_i+\hat{q}_i)} \text{ and } 1 \leqslant w \leqslant \min_i \frac{2+w'(1-\hat{q}_i)}{2+\hat{b}_i},$$

$$i \in \langle n \rangle, \quad (3.5)$$

if $w > w'$, *and for*

$$0 < w' < \min_i \frac{2-\hat{b}_i}{1+\hat{q}_i} \text{ and } \max_i \frac{w'(1+\hat{q}_i)}{2-\hat{b}_i} < w \leqslant w', \qquad i \in \langle n \rangle, \quad (3.6)$$

or

$$\frac{2}{1+\min_i(\hat{q}_i+\hat{b}_i)} < w' < \min_i \frac{2-\hat{b}_i}{1+\hat{q}_i} \text{ and } 1 \leqslant w < \min_i \frac{2-w'(1+\hat{q}_i)}{\hat{b}_i},$$

$$i \in \langle n \rangle, \quad (3.7)$$

if $w \leqslant w'$.

THEOREM 5. *If A of* (1.2) *is irreducible weakly diagonally dominant, then the MSOR method is convergent for* w *and* w' *given by* (3.4), (3.5), (3.6), (3.7).

Proof. As is known, an irreducible weakly diagonally dominant matrix can be transformed by scaling in a strictly diagonally dominant matrix. So it is generalized diagonally dominant by rows and columns (see [10]). Thus this result comes from the last corollary. ∎

THEOREM 6. *If $A = I - Q - B$ is an M, H, or Stieltjes matrix, then the MSOR method is convergent for w and w' given by* (3.4), (3.5), (3.6), (3.7).

Proof. It follows, by Corollary 2 and Theorem 5 of [2], from Theorem 4 of [3] and Theorem 1 of [9]. ■

4. VECTORIAL NORMS AND THEIR APPLICATION TO THE STUDY OF THE MSOR METHOD'S CONVERGENCE

Let G be a linear operator defined over $\mathbb{C}^n$, and let us consider the following partition of $\mathbb{C}^n$ as a direct sum:

$$\mathbb{C}^n = W_1 \oplus W_2 \oplus \cdots \oplus W_k.$$

Let P_i be a projection operator of $x \in \mathbb{C}^n$ over W_i, where $x_i = P_i x$ denotes the projection of $x \in \mathbb{C}^n$ onto W_i.

Let us denote by ϕ_i a norm over W_i.

DEFINITION 9. Let p be a map of $\mathbb{C}^n$ in $\mathbb{R}^k_+$. Then p is a vectorial regular norm of dimension k over $\mathbb{C}^n$ if it satisfies

$$x \in \mathbb{C}^n \quad \to \quad p(x) = \begin{bmatrix} \phi_1 P_1(x) \\ \vdots \\ \phi_k P_k(x) \end{bmatrix} = \begin{bmatrix} \phi_1(x_1) \\ \vdots \\ \phi_k(x_k) \end{bmatrix} \in \mathbb{R}^k_+.$$

DEFINITION 10. The majorant of G is the matrix $M(k,k)$ whose elements are real and that verifies

$$\forall x \in \mathbb{C}^n \qquad p(Gx) \leqslant Mp(x).$$

THEOREM 7 [6]. *The matrix $M(G) = \{m_{ij}(G)\}$ defined by*

$$m_{ij}(G) = \sup_{\substack{x_j \in W_j \\ x_j \neq 0}} \left\{ \frac{\phi_i(G_{ij}x_j)}{\phi_j(x_j)} \right\} \underset{\text{def}}{=} S_{\phi_i \phi_j}(G_{ij})$$

is the smallest of the majorants of A. The map $G \to M(G)$ is a vectorial regular submultiplicative norm of order k^2 over $L(\mathbb{C}^n, \mathbb{C}^n)$.

DEFINITION 11. G is said to be contractive in relation to the vectorial norm p if it has, relative to p, a majorant which is convergent.

Thus, to say that G is contractive relatively to p is to say that

$$\rho(M(G)) < 1.$$

Let us now present some theorems and a lemma, which will be useful in the proof of the most important theorem of this section.

THEOREM 8 [6]. *All the majorants of G, particularly $M(G)$ satisfy*:

$$\forall G \in L(\mathbb{C}^n, \mathbb{C}^n) \qquad \rho(G) \leqslant \rho(M(G)).$$

LEMMA 2 [6]. *If T is contractive relative to p, then $(I-T)^{-1}$ exists and for all convergent majorants B of T we have*

$$M(I-T)^{-1} \leqslant [I - M(T)]^{-1} \leqslant (I-B)^{-1}.$$

THEOREM 9 [8]. *If T is contractive relatively to p, it is convergent. The linear iteration*

$$x_{r+1} = Tx_r + h$$

converges to the unique solution $x = (I-T)^{-1}h$ of the linear system $(I-T)x = h$. Thus if we fix an interval r_0, for all $r \geqslant r_0$ and for all convergent majorants B (of T) we have the following upper bounds for the error expressed in terms of vectorial norms:

$$\begin{aligned} p(\varepsilon_r) \underset{\text{def}}{=} p(x_r - x) &\leqslant [M(T)]^{r-r_0}[I - M(T)]^{-1} p(x_{r_0+1} - x_{r_0}) \\ &\leqslant B^{r-r_0}[I-B]^{-1} p(x_{r_0+1} - x_{r_0}) \end{aligned}$$

Now, we present the Stein-Rosenberg theorem in a particular formulation, and then prove the very useful Theorem 10.

DEFINITION 12. A is said to be a block-H matrix relatively to a vectorial norm p if the correspondent Jacobi matrix for blocks is contracted relative to p.

STEIN-ROSENBERG THEOREM [6]. *Let U and V be two nonnegative matrices. The two following propositions are equivalent*:

(1) $\rho(U+V)<1$,
(2) $\rho(U)<1$ *and* $\rho([I-U]^{-1}V)<1$.

If they are verified, then we have

$$\rho\left([I-U]^{-1}V\right)\leqslant\rho(U+V)<1.$$

THEOREM 10. *Let p be a vectorial regular norm of dimension k over $\mathbb{R}^n$, and M a vectorial matricial norm generated by p. Let Q and R be two $n\times n$ matrices. If $\rho=\rho(M(Q)+M(B))$ verifies $\rho<1$, then the iteration matrix of the MSOR method,*

$$L_{w,w'}=(I-w'Q)^{-1}[(1-w)I+wB+(w-w')R],$$

exists, and it is contractive relative to p for values of w and w' such that

$$0<w'\leqslant w<1 \quad or \quad 1\leqslant w\leqslant w'<\frac{2}{1+\rho}.$$

This condition defines the set of the values w and w' such that the matrix

$$\overline{L}_{w,w'}=[I-|w'|M(Q)]^{-1}[|1-w|I+|w|M(B)+|w-w'|R]$$

is a convergent majorant of $L_{w,w'}$.

Proof. First we shall prove that the matrix $w'Q$ is contractive if $\rho<1$ and $0<w'<2/(1+\rho)$. In fact,

$$\rho\left(M(w'Q)\right)\leqslant\rho\left(M(w'B)+M(w'Q)\right)=|w'|\rho<\frac{2\rho}{1+\rho}<1.$$

If we use the last lemma, we can say:

(a) $(I-w'Q)^{-1}$ exists and we have the matrix $L_{w,w'}$.
(b) $M(I-w'Q)^{-1}\leqslant[I-M(w'Q)]^{-1}=[I-|w'|M(Q)]^{-1}$.

Thus

$$M(L_{w,w'}) = M\{(I - w'Q)^{-1}[(1-w)I + (w - w')R + wB]\}$$
$$\leqslant M(I - w'Q)^{-1}M[(1-w)I + (w-w')R + wB]$$
$$\leqslant [I - |w'|M(Q)]^{-1}[|1-w|I + |w-w'|R + |w|M(B)].$$

Then

$$\bar{L}_{w,w'} = [I - |w'|M(Q)]^{-1}[(1-w)I + |w-w'|R + |w|M(B)]$$

is a majorant of $L_{w,w'}$ relative to p.

Let us define

$$S = M(w'Q) = |w'|M(Q),$$
$$T = |1-w|I + |w-w'|R + |w|M(B),$$

By the Stein-Rosenberg theorem above, we can say the $L_{w,w'}$ is convergent if and only if $\rho(S+T) < 1$, that is,

$$\rho[|w'|M(Q) + |1-w|I + |w-w'|R + |w|M(B)] < 1.$$

Let us prove this last relation for $0 < w \leqslant w' < 1$ and for $1 \leqslant w \leqslant w' < 2/(1+\rho)$. We consider two cases:

Case 1: $1 \leqslant w \leqslant w' < 2/(1+\rho)$. For these values of w and w', we have

$$\rho[|w'|M(Q) + |1-w|I + |w-w'|R + |w|M(B)]$$
$$\leqslant \rho\{|w'|[M(Q) + M(B)] + |1-w|I + |w-w'|R\}$$
$$\leqslant |w'|\rho + \max\{w-1, w'-1\}$$
$$= w'\rho + w' - 1.$$

Then, as $w'\rho + w' - 1 < 1$ for $0 < w' < 2/(1+\rho)$, it is evident that $L_{w,w'}$ is contractive relative to p, for $1 \leqslant w \leqslant w' < 2/(1+\rho)$. So by Theorem 9 we can say that the MSOR method is convergent for $1 \leqslant w \leqslant w' < 2/(1+\rho)$ if $\rho < 1$.

Case 2: $0 < w' \leqslant w < 1$. For these values, we have

$$\rho\{|w'|M(Q) + |1-w|I + |w-w'|R + |w|M(B)\}$$
$$\leqslant \rho\{|w'|[M(Q)+M(B)] + (1-w)I + (w-w')R\}$$
$$\leqslant w'\rho + \max\{(1-w),(1-w')\}$$
$$= w'\rho + 1 - w'.$$

Then

$$[|w'|M(Q) + |1-w|I + |w-w'|R + |w|M(B)] < 1$$

for $0 < w' \leqslant w < 1$ and $\rho < 1$. In these conditions, we conclude that $\bar{L}_{w,w'}$ is a convergent majorant of $L_{w,w'}$; then $L_{w,w'}$ is contractive relative to p, and the MSOR method is convergent. ■

Since for $w = w'$ the MSOR method is the well-known SOR method, we can now present Theorem 6 of [6] as a corollary of this theorem when A of (1.2) has property A.

COROLLARY 3 [6]. *Let p be a vectorial norm of dimension k over $\mathbb{R}^n$, and let M be a vectorial matricial norm generated by p and $\rho = \rho(M(Q) + M(B)) < 1$. Then*:

(1) *$J = Q + B$ and $L_1 = (I-Q)^{-1}B$ are contractive relative to p.*

(2) *The matrix of the overrelaxation method, $L_w = (I - wQ)^{-1}[wB + (1-w)I]$, exists and is contractive for all w's such that*

$$0 < w < \frac{2}{1+\rho}.$$

This condition defines the set of values of w such that the matrix

$$\bar{L}_w \underset{\text{def}}{=} [I - |w|M(Q)]^{-1}[(1-w)I + |w|B]$$

is a convergent majorant of L_w.

Now, if we consider the vectorial norm $p(x) = |x|$ and the iterative methods by points [all the blocks are (1,1)], we have the following result.

THEOREM 11. *If $A = I - Q - B$ is an H-matrix, then the MSOR method is convergent for w and w' given by any of the conditions*:

(i) $0 < w' \leqslant w < 1$, *if* $\min_i w'(1-\hat{q}_i)/\hat{b}_i < 1$ *with* $0 < w' < \min_i \hat{b}_i/(1-\hat{q}_i)$, $i \in \langle n \rangle$.

(ii) $0 < w' < \min_i \hat{b}_i/(1-\hat{q}_i)$ *and* $w' \leqslant w < \min_i w'(1-\hat{q}_i)/\hat{b}_i$, $i \in \langle n \rangle$.

(iii) $\min_i \hat{b}_i/(1-\hat{q}_i) < w' < 1$ *and* $\min_i w'(1-\hat{q}_i)/\hat{b}_i \leqslant w < 1$ *with* $w' \leqslant w$, $i \in \langle n \rangle$.

(iv) $\max \hat{b}_i/(1-\hat{q}_i) < w' < 2/[1+\max_i(\hat{b}_i+\hat{q}_i)]$ *and* $1 \leqslant w \leqslant \min_i [2+w'(1-\hat{q}_i)]/\hat{b}_i$, $i \in \langle n \rangle$, *if* $w > w'$.

(v) $0 < w' < \min_i (2-\hat{b}_i)/(1+\hat{q}_i)$ *and* $\max w'(1+\hat{q}_i)/(2-\hat{b}_i) < w < w'$, $i \in \langle n \rangle$.

(vi) $2/[1+\min_i(\hat{q}_i+\hat{b}_i)] < w' < \min_i (2-\hat{b}_i)/(1+\hat{q}_i)$ *and* $1 \leqslant w < \min_i [2-w'(1+\hat{q}_i)]/\hat{b}_i$, $i \in \langle n \rangle$, *if* $w \leqslant w'$.

(vii) $1 \leqslant w \leqslant w' \leqslant 2/(1+\rho)$.

Proof. The intervals of convergence obtained in this theorem are the joins of the intervals achieved by Corollary 2 and Theorem 10. ∎

In this theorem we have given the largest known intervals of convergence for the MSOR method, when $A = I - Q - B$ is an H-matrix. This result is very general, since by [9] we know that the class of H-matrices involves many types of matrices.

The author wishes to thank Professor Fernanda de Oliveira and the referees of this paper for their suggestions and criticism during the writing of this work.

REFERENCES

1 G. H. Golub and R. S. Varga, Chebyshev semi-iterative methods, successive overrelaxation iterative methods, and second-order Richardson iterative methods, Parts I and II, *Numer. Math.* 3:147–168 (1961).
2 K. R. James and A. Riha, Convergence criteria for successive overrelaxation, *SIAM J. Numer. Anal.* 12(2):137–143 (1975).
3 M. Martins, Generalized diagonal dominance in connection with the accelerated overrelaxation (AOR) method, *BIT* 22:73–78 (1982).
4 M. Martins, On accelerated overrelaxation iterative method for linear systems with strictly diagonally dominant matrices, *Math. Comp.* 35:1269–1273 (1980).
5 M. Martins, An improvement for the area of convergence of the accelerated overrelaxation iterative method, *Rev. Anal. Numér. Théorie Approximation* 12:65–76 (1983).

6 F. Robert, Block *H*-matrices et convergence des methodes iteratives classiques par blocs, *Linear Algebra Appl.* 2:223–265 (1969).
7 J. W. Sheldon, On the spectral norms of several iterative processes, *J. Assoc. Comput. Mach.* 6:494–505 (1959).
8 J. Schroeder, Computing error bounds in solving linear systems, MRC Tech. Report 242, Univ. of Wisconsin, July 1961.
9 R. S. Varga, On recurring theorems on diagonal dominance, *Linear Algebra Appl.* 13:1–9 (1976).
10 W. Walter, Bemerkungen zu Iterationsverfahren bei linearen Gleichungs-systemen, *Numer. Math.* 10:80–85 (1976).
11 D. M. Young, *Iterative Solution of Large Linear Systems*, Academic, New York, 1971.
12 D. M. Young and H. E. Eidson, On the determination of the optimum relaxation factor for the SOR method when the eigenvalues of the Jacobi method are complex, Report CNA-1, Center for Numerical Analysis, Univ. of Texas, Austin, 1969.

Received February 1983; revised 7 September 1985